

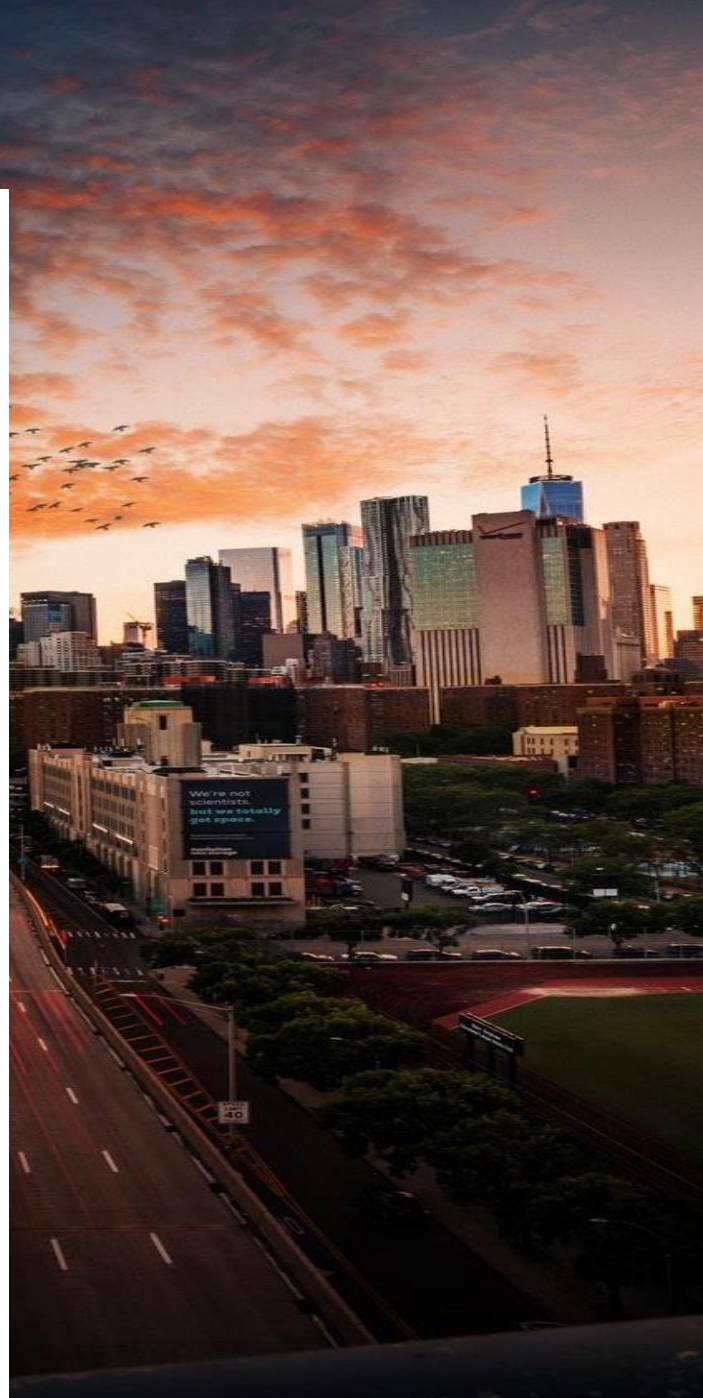
# Coursera Capstone 2019

---

---

**Opening a New Restaurant in Mumbai, India**

**Authored by: Reshma Yemulwar**



---

# Opening a New Restaurant in Mumbai, India

## Introduction

For many shoppers, visiting restaurants is a great way to relax and enjoy themselves during weekends and holidays. They can do grocery shopping, dine at restaurants, shop at the various fashion outlets, watch movies and perform many more activities. Restaurants are like a one-stop destination for all types of food lovers. For retailers, the central location and the large crowd at the restaurants provides a great distribution channel to market their products and services. Property developers are also taking advantage of this trend to build more restaurants to cater to the demand. As a result, there are many restaurants in Mumbai and many more are being built. Opening restaurants allows property developers to earn consistent rental income. Of course, as with any business decision, opening a new restaurant requires serious consideration and is a lot more complicated than it seems. Particularly, the location of the restaurant is one of the most important decisions that will determine whether the restaurant will be a success or a failure.

## Business Problem

The objective of this capstone project is to analyze and select the best locations in Mumbai, India in order to open a new restaurant. Using data science methodology and machine learning techniques like clustering, this project aims to provide solutions to answer the business question: In Mumbai, India, if a property developer is looking to open a new restaurant, where would you recommend that they open it?

---

## Target Audience for this Project

This project is particularly useful to property developers and investors looking to open or invest in new restaurants in the financial capital city of India i.e. Mumbai

## Data

**To solve the problem, we will need the following data:**

- List of neighborhoods in Mumbai. This defines the scope of this project which is confined to the city of Mumbai, the financial capital city of India in Asia.
- Latitude and longitude coordinates of those neighborhoods. This is required in order to plot the map and also to get the venue data.
- Venue data, particularly data related to restaurants. We will use this data to perform clustering on the neighborhoods.

**Sources of data and methods to extract them:**

This Wikipedia page ([https://en.wikipedia.org/wiki/List\\_of\\_neighbourhoods\\_in\\_Mumbai](https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Mumbai)) contains a list of neighborhoods in Mumbai, with a total of 37 neighborhoods. We will use web scraping techniques to extract the data from the Wikipedia page, with the help of Python requests and BeautifulSoup packages. Then we will get the geographical coordinates of the neighborhoods using Python Geocoder package which will give us the latitude and longitude coordinates of the neighborhoods.

After that, we will use Foursquare API to get the venue data for those neighborhoods. Foursquare has one of the largest database of 105+ million places and is used by over

---

125,000 developers. Foursquare API will provide many categories of the venue data, we are particularly interested in the Restaurant category in order to help us to solve the business problem put forward. This is a project that will make use of many data science skills, from web scraping (Wikipedia), working with API (Foursquare), data cleaning, data wrangling, to machine learning (K-means clustering) and map visualization (Folium). In the next section, we will present the Methodology section where we will discuss the steps taken in this project, the data analysis that we did and the machine learning technique that was used.